

What is claimed is:

1 1. A method for receiving a video sequence including query
2 objects to be extracted and generating object-labeled images based on the
3 query objects, the method comprising the steps of:

4 (a) dividing the video sequence into one or more shots, each of which
5 is a set of frames having a similar scene, and selecting one or more key
6 frames from each of the shots;

7 (b) extracting query object based initial object regions from each of
8 the key frames;

9 (c) tracking object regions in all frames of each of the shots based on
10 the corresponding query image based initial object regions; and

11 (d) labeling the object regions tracked in each of the frames based on
12 information on the corresponding query objects.

1 2. The method of claim 1, wherein step (b) comprises:

2 (b1) determining whether there exists an object similar to each of the query
3 objects in each of the key frames, and if there is a similar object in a key
4 frame, extracting the similar object as a corresponding query object based
5 initial object region; and

6 (b2) generating query object based shot mask images in all key
7 frames of the shots by setting pixels of the query object based initial object
8 regions extracted from each of the key frames as a first value and setting
9 the remaining pixels of each of the key frames as a second value.

1 3. The method of claim 2, wherein step (c) comprises:

2 (c1) tracking the object regions in all frames of each of the shots based on
3 the corresponding query image based shot mask images and video feature
4 values of the corresponding query objects; and

5 (c2) generating query object based frame mask images in all frames

6 all key frames of each of the shots by setting pixels of the query object
7 based initial object regions extracted from each of the key frames as a first
8 value and setting the remaining pixels of each of the key frames as a
9 second value.

1 7. The apparatus of claim 6, wherein the object region tracker
2 tracks the object regions in all frames of each of the shots based on the
3 corresponding query image based shot mask images and video feature
4 values of the corresponding query objects, and generates query object
5 based frame mask images in all frames of each of the shots by setting pixels
6 of the object regions tracked in each of the frames as a first value and
7 setting the remaining pixels of each of the key frames as a second value.

1 8. The apparatus of claim 5, wherein the object-labeled image
2 generator labels each of the object regions in each frame with a unique
3 number set to the corresponding query image or coordinate information of
4 the corresponding query image in each frame.

1 9. A computer readable medium having embodied thereon a
2 computer program for receiving a video sequence including query objects to
3 be extracted and generating object-labeled images based on the query
4 objects, wherein generating object-labeled images comprises the steps of:
5 (a) dividing the video sequence into one or more shots, each of which
6 is a set of frames having a similar scene, and selecting one or more key
7 frames from each of the shots;
8 (b) extracting query object based initial object regions from each of
9 the key frames;
10 (c) tracking object regions in all frames of each of the shots based on
11 the corresponding query image based initial object regions; and

- 12 (d) labeling the object regions tracked in each of the frames based on
13 information on the corresponding query objects.